

RECEIVED



FEB 21 2002

OIPE

TECH CENTER 1600/2900

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/09/981,947**

DATE: 02/07/2002  
TIME: 15:16:24

Input Set : N:\Crf3\RULE60\09981947.raw  
Output Set: N:\CRF3\02072002\I981947.raw

## **SEQUENCE LISTING**

**ENTERED**

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/981,947

DATE: 02/07/2002  
TIME: 15:16:24

Input Set : N:\CrF3\RULE60\09981947.raw  
Output Set: N:\CRF3\02072002\I981947.raw

45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
46	TCGACCCACG CGTCCGGCCT TGGCAGAGTC TGGGGTCCCT GGACTGAGCC ATCAGCTGGG	60
47	TCACTGAGAC CC ATG GCA AGG AAA CAA AAT AGG AAT TCC AAG GAA CTG GGC	111
48	Met Ala Arg Lys Gln Asn Arg Asn Ser Lys Glu Leu Gly	
49	1 5 10	
50	CTA GTT CCC CTC ACA GAT GAC ACC AGC CAC GCC GGG CCT CCA GGG CCA	159
51	Leu Val Pro Leu Thr Asp Asp Thr Ser His Ala Gly Pro Pro Gly Pro	
52	15 20 25	
53	Gly Arg Ala Leu Leu Glu Cys Asp His Leu Arg Ser Gly Val Pro Gly	207
54	30 35 40 45	
55	GGA AGG AGA AGA AAG GAC TGG TCC TGC TCG CTC CTC GTG GCC TCC CTC	255
56	Gly Arg Arg Arg Lys Asp Trp Ser Cys Ser Leu Leu Val Ala Ser Leu	
57	50 55 60	
58	GCG GGC GCC TTC GGC TCC TCC TTC CTC TAC GGC TAC AAC CTG TCG GTG	303
59	Ala Gly Ala Phe Gly Ser Ser Phe Leu Tyr Gly Tyr Asn Leu Ser Val	
60	65 70 75	
61	GTG AAT GCC CCC ACC CCG TAC ATC AAG GCC TTT TAC AAT GAG TCA TGG	351
62	Val Asn Ala Pro Thr Pro Tyr Ile Lys Ala Phe Tyr Asn Glu Ser Trp	
63	80 85 90	
64	GAA AGA AGG CAT GGA CGT CCA ATA GAC CCA GAC ACT CTG ACT CTG CTC	399
65	Glu Arg Arg His Gly Arg Pro Ile Asp Pro Asp Thr Leu Thr Leu	
66	95 100 105	
67	TGG TCT GTG ACT GTG TCC ATA TTC GCC ATC GGT GGA CTT GTG GGG ACG	447
68	Trp Ser Val Thr Val Ser Ile Phe Ala Ile Gly Gly Leu Val Gly Thr	
69	110 115 120 125	
70	TTA ATT GTG AAG ATG ATT GGA AAG GTT CTT GGG AGG AAG CAC ACT TTG	495
71	Leu Ile Val Lys Met Ile Gly Lys Val Leu Gly Arg Lys His Thr Leu	
72	130 135 140	
73	CTG GCC AAT AAT GGG TTT GCA ATT TCT GCT GCA TTG CTG ATG GCC TGC	543
74	Leu Ala Asn Asn Gly Phe Ala Ile Ser Ala Ala Leu Leu Met Ala Cys	
75	145 150 155	
76	TCG CTC CAG GCA GGA GCC TTT GAA ATG CTC ATT GTG GGA CGC TTC ATC	591
77	Ser Leu Gln Ala Gly Ala Phe Glu Met Leu Ile Val Gly Arg Phe Ile	
78	160 165 170	
79	ATG GGC ATA GAT GGA GGC GTC GCC CTC AGT GTG CTC CCC ATG TAC CTC	639
80	Met Gly Ile Asp Gly Val Ala Leu Ser Val Leu Pro Met Tyr Leu	
81	175 180 185	
82	AGT GAG ATC TCA CCC AAG GAG ATC CGT GGC TCT CTG GGG CAG GTG ACT	687
83	Ser Glu Ile Ser Pro Lys Glu Ile Arg Gly Ser Leu Gly Gln Val Thr	
84	190 195 200 205	
85	GCC ATC TTT ATC TGC ATT GGC GTG TTC ACT GGG CAG CTT CTG GGC CTG	735
86	Ala Ile Phe Ile Cys Ile Gly Val Phe Thr Gly Gln Leu Leu Gly Leu	
87	210 215 220	
88	CCC GAG CTG CTG GGA AAG GAG AGT ACC TGG CCA TAC CTG TTT GGA GTG	783
89	Pro Glu Leu Leu Gly Lys Glu Ser Thr Trp Pro Tyr Leu Phe Gly Val	
90	225 230 235	
91	ATT GTG GTC CCT GCC GTT GTC CAG CTG CTG AGC CTT CCC TTT CTC CCG	831
92	Ile Val Val Pro Ala Val Val Gln Leu Leu Ser Leu Pro Phe Leu Pro	
93		

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/981,947

DATE: 02/07/2002  
TIME: 15:16:24

Input Set.: N:\CRF3\RULE60\09981947.raw  
Output Set: N:\CRF3\02072002\I981947.raw

94	240	245	250	
95	GAC AGC CCA CGC TAC CTG CTC TTG GAG AAG CAC AAC GAG GCA AGA GCT			879
96	Asp Ser Pro Arg Tyr Leu Leu Leu Glu Lys His Asn Glu Ala Arg Ala			
97	255	260	265	
98	GTG AAA GCC TTC CAA ACG TTC TTG GGT AAA GCA GAC GTT TCC CAA GAG			927
99	Val Lys Ala Phe Gln Thr Phe Leu Gly Lys Ala Asp Val Ser Gln Glu			
100	270	275	280	285
101	GTA GAG GAG GTC CTG GCT GAG AGC CAC GTG CAG AGG AGC ATC CGC CTG			975
102	Val Glu Glu Val Leu Ala Glu Ser His Val Gln Arg Ser Ile Arg Leu			
103	290	295	300	
104	GTC TCC GTG CTG GAG CTG AGA GCT CCC TAC GTC CGC TGG CAG GTG			1023
105	Val Ser Val Leu Glu Leu Leu Arg Ala Pro Tyr Val Arg Trp Gln Val			
106	305	310	315	
107	GTC ACC GTG ATT GTC ACC ATG GCC TGC TAC CAG CTC TGT GGC CTC AAT			1071
108	Val Thr Val Ile Val Thr Met Ala Cys Tyr Gln Leu Cys Gly Leu Asn			
109	320	325	330	
110	GCA ATT TGG TTC TAT ACC AAC AGC ATC TTT GGA AAA GCT GGG ATC CCT			1119
111	Ala Ile Trp Phe Tyr Thr Asn Ser Ile Phe Gly Lys Ala Gly Ile Pro			
112	335	340	345	
113	CCG GCA AAG ATC CCA TAC GTC ACC TTG AGT ACA GGG GGC ATC GAG ACT			1167
114	Pro Ala Lys Ile Pro Tyr Val Thr Leu Ser Thr Gly Gly Ile Glu Thr			
115	350	355	360	365
116	TTG GCT GCC GTC TTC TCT GGT TTG GTC ATT GAG CAC CTG GGA CGG AGA			1215
117	Leu Ala Ala Val Phe Ser Gly Leu Val Ile Glu His Leu Gly Arg Arg			
118	370	375	380	
119	CCC CTC CTC ATT GGT GGC TTT GGG CTC ATG GGC CTC TTC TTT GGG ACC			1263
120	Pro Leu Leu Ile Gly Gly Phe Gly Leu Met Gly Leu Phe Phe Gly Thr			
121	385	390	395	
122	CTC ACC ATC ACG CTG ACC CTG CAG GAC CAC GCC CCC TGG GTC CCC TAC			1311
123	Leu Thr Ile Thr Leu Thr Leu Gln Asp His Ala Pro Trp Val Pro Tyr			
124	400	405	410	
125	CTG AGT ATC GTG GGC ATT CTG GCC ATC ATC GCC TCT TTC TGC AGT GGG			1359
126	Leu Ser Ile Val Gly Ile Leu Ala Ile Ile Ala Ser Phe Cys Ser Gly			
127	415	420	425	
128	CCA GGT GGC ATC CCG TTC ATC TTG ACT GGT GAG TTC TTC CAG CAA TCT			1407
129	Pro Gly Gly Ile Pro Phe Ile Leu Thr Gly Glu Phe Phe Gln Gln Ser			
130	430	435	440	445
131	CAG CGG CCG GCT GCC TTC ATC ATT GCA GGC ACC GTC AAC TGG CTC TCC			1455
132	Gln Arg Pro Ala Ala Phe Ile Ile Ala Gly Thr Val Asn Trp Leu Ser			
133	450	455	460	
134	AAC TTT GCT GTT GGG CTC CTC TTC CCA TTC ATT CAG AAA AGT CTG GAC			1503
135	Asn Phe Ala Val Gly Leu Leu Phe Pro Phe Ile Gln Lys Ser Leu Asp			
136	465	470	475	
137	ACC TAC TGT TTC CTA GTC TTT GCT ACA ATT TGT ATC ACA GGT GCT ATC			1551
138	Thr Tyr Cys Phe Leu Val Phe Ala Thr Ile Cys Ile Thr Gly Ala Ile			
139	480	485	490	
140	TAC CTG TAT TTT GTG CTG CCT GAG ACC AAA AAC AGA ACC TAT GCA GAA			1599
141	Tyr Leu Tyr Phe Val Leu Pro Glu Thr Lys Asn Arg Thr Tyr Ala Glu			
142	495	500	505	

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/981,947

DATE: 02/07/2002  
TIME: 15:16:24

Input Set : N:\Crf3\RULE60\09981947.raw  
Output Set: N:\CRF3\02072002\I981947.raw

143	ATC AGC CAG GCA TTT TCC AAA AGG AAC AAA GCA TAC CCA CCA GAA GAG	1647
144	Ile Ser Gln Ala Phe Ser Lys Arg Asn Lys Ala Tyr Pro Pro Glu Glu	
145	510 515 520 525	
146	AAA ATC GAC TCA GCT GTC ACT GAT GCT CCT GCT TCT TCT CCT TTC ACT	1695
147	Lys Ile Asp Ser Ala Val Thr Asp Ala Pro Ala Ser Ser Pro Phe Thr	
148	530 535 540	
149	ACT CCG AAT ACA GCC TGG ATT CAA GCT GCC GCC ACC ACC ACC GCC ACC	1743
150	Thr Pro Asn Thr Ala Trp Ile Gln Ala Ala Ala Thr Thr Thr Ala Thr	
151	545 550 555	
152	AAA AAA GAA CAC CCA TTG TAAACGGTCA TGTGGTATT CCTCAACCTG GAATGACC	1799
153	Lys Lys Glu His Pro Leu	
154	560	
155	TTCCCCATTC TTCTTCTCCT GGAGAACACC AAGTCATGAT GTCAGACAAG AGCTTGGATT	1859
156	TTGGAGACAT GGGTTGAAT TCCAGTCATT CATTCTTTA TTCAGCAAAT ATTTAACAAAG	1919
157	TACTGACATG TCCCATAATGT TGTTTACCC ACTGGTTATA CAATGGGAGG GAGAGAGAGA	1979
158	GAGAGAGAGA GAGAGAGATG CTATTCTAA AGCTTGAAGT CTAGGCTGTG CACGGTGGCT	2039
159	CACGCCTGTA ATCCCAGCAC TTGGGGAGGC CGAGGTGGGT GGATCGTGAG GTCAGGAGAT	2099
160	TGAGACCATC CTGGCTAACAA TGTTGAAACT CCCTCTCTAC TAAAAATACA AAAAATTAGC	2159
161	TGAGCATGGT GGCGGGCGCC TGAGTCCCA GCTACTTGGG AGGCTGAGGC AGGAGAATGG	2219
162	CGTGAACCCA GGAGGCGGAG CTTGCAGTGA GCCGAGATCA CACCACCACA CTCCAGCCTG	2279
163	GGTGACAGAG CCAGACTCCG TCTCAAAAAA AAAAAAAA AAAAGGGCGG	2339
164	CCGC	2343

166 (2) INFORMATION FOR SEQ ID NO: 2:

167 (i) SEQUENCE CHARACTERISTICS:  
 168 (A) LENGTH: 563 amino acids  
 169 (B) TYPE: amino acid  
 170 (D) TOPOLOGY: linear

171 (ii) MOLECULE TYPE: protein  
 172 (v) FRAGMENT TYPE: internal  
 173 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

174	Met Ala Arg Lys Gln Asn Arg Asn Ser Lys Glu Leu Gly Leu Val Pro	
175	1 5 10 15	
176	Leu Thr Asp Asp Thr Ser His Ala Gly Pro Pro Gly Pro Gly Arg Ala	
177	20 25 30	
178	Leu Leu Glu Cys Asp His Leu Arg Ser Gly Val Pro Gly Gly Arg Arg	
179	35 40 45	
180	Arg Lys Asp Trp Ser Cys Ser Leu Leu Val Ala Ser Leu Ala Gly Ala	
181	50 55 60	
182	Phe Gly Ser Ser Phe Leu Tyr Gly Tyr Asn Leu Ser Val Val Asn Ala	
183	65 70 75 80	
184	Pro Thr Pro Tyr Ile Lys Ala Phe Tyr Asn Glu Ser Trp Glu Arg Arg	
185	85 90 95	
186	His Gly Arg Pro Ile Asp Pro Asp Thr Leu Thr Leu Leu Trp Ser Val	
187	100 105 110	
188	Thr Val Ser Ile Phe Ala Ile Gly Gly Leu Val Gly Thr Leu Ile Val	
189	115 120 125	
190	Lys Met Ile Gly Lys Val Leu Gly Arg Lys His Thr Leu Leu Ala Asn	
191	130 135 140	
192	Asn Gly Phe Ala Ile Ser Ala Ala Leu Leu Met Ala Cys Ser Leu Gln	

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/981,947

DATE: 02/07/2002  
TIME: 15:16:24

Input Set : N:\Crf3\RULE60\09981947.raw  
Output Set: N:\CRF3\02072002\I981947.raw

193	145	150	155	160													
194	Ala	Gly	Ala	Phe	Glu	Met	Leu	Ile	Val	Gly	Arg	Phe	Ile	Met	Gly	Ile	
195						165				170					175		
196	Asp	Gly	Gly	Val	Ala	Leu	Ser	Val	Leu	Pro	Met	Tyr	Leu	Ser	Glu	Ile	
197						180				185					190		
198	Ser	Pro	Lys	Glu	Ile	Arg	Gly	Ser	Leu	Gly	Gln	Val	Thr	Ala	Ile	Phe	
199						195				200					205		
200	Ile	Cys	Ile	Gly	Val	Phe	Thr	Gly	Gln	Leu	Leu	Gly	Leu	Pro	Glu	Leu	
201						210				215					220		
202	Leu	Gly	Lys	Glu	Ser	Thr	Trp	Pro	Tyr	Leu	Phe	Gly	Val	Ile	Val	Val	
203						225				230					240		
204	Pro	Ala	Val	Val	Gln	Leu	Leu	Ser	Leu	Pro	Phe	Leu	Pro	Asp	Ser	Pro	
205						245				250					255		
206	Arg	Tyr	Leu	Leu	Glu	Lys	His	Asn	Glu	Ala	Arg	Ala	Val	Lys	Ala		
207						260				265					270		
208	Phe	Gln	Thr	Phe	Leu	Gly	Lys	Ala	Asp	Val	Ser	Gln	Glu	Val	Glu	Glu	
209						275				280					285		
210	Val	Leu	Ala	Glu	Ser	His	Val	Gln	Arg	Ser	Ile	Arg	Leu	Val	Ser	Val	
211						290				295					300		
212	Leu	Glu	Leu	Leu	Arg	Ala	Pro	Tyr	Val	Arg	Trp	Gln	Val	Val	Thr	Val	
213						305				310					320		
214	Ile	Val	Thr	Met	Ala	Cys	Tyr	Gln	Leu	Cys	Gly	Leu	Asn	Ala	Ile	Trp	
215						325				330					335		
216	Phe	Tyr	Thr	Asn	Ser	Ile	Phe	Gly	Lys	Ala	Gly	Ile	Pro	Pro	Ala	Lys	
217						340				345					350		
218	Ile	Pro	Tyr	Val	Thr	Leu	Ser	Thr	Gly	Gly	Ile	Glu	Thr	Leu	Ala	Ala	
219						355				360					365		
220	Val	Phe	Ser	Gly	Leu	Val	Ile	Glu	His	Leu	Gly	Arg	Arg	Pro	Leu	Leu	
221						370				375					380		
222	Ile	Gly	Gly	Phe	Gly	Leu	Met	Gly	Leu	Phe	Phe	Gly	Thr	Leu	Thr	Ile	
223						385				390					400		
224	Thr	Leu	Thr	Leu	Gln	Asp	His	Ala	Pro	Trp	Val	Pro	Tyr	Leu	Ser	Ile	
225						405				410					415		
226	Val	Gly	Ile	Leu	Ala	Ile	Ile	Ala	Ser	Phe	Cys	Ser	Gly	Pro	Gly	Gly	
227						420				425					430		
228	Ile	Pro	Phe	Ile	Leu	Thr	Gly	Glu	Phe	Phe	Gln	Gln	Ser	Gln	Arg	Pro	
229						435				440					445		
230	Ala	Ala	Phe	Ile	Ile	Ala	Gly	Thr	Val	Asn	Trp	Leu	Ser	Asn	Phe	Ala	
231						450				455					460		
232	Val	Gly	Leu	Leu	Phe	Pro	Phe	Ile	Gln	Lys	Ser	Leu	Asp	Thr	Tyr	Cys	
233						465				470					475		
234	Phe	Leu	Val	Phe	Ala	Thr	Ile	Cys	Ile	Thr	Gly	Ala	Ile	Tyr	Leu	Tyr	
235						485				490					495		
236	Phe	Val	Leu	Pro	Glu	Thr	Lys	Asn	Arg	Thr	Tyr	Ala	Glu	Ile	Ser	Gln	
237						500				505					510		
238	Ala	Phe	Ser	Lys	Arg	Asn	Lys	Ala	Tyr	Pro	Pro	Glu	Glu	Lys	Ile	Asp	
239						515				520					525		
240	Ser	Ala	Val	Thr	Asp	Ala	Pro	Ala	Ser	Ser	Pro	Phe	Thr	Thr	Pro	Asn	
241						530				535					540		

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/981,947

DATE: 02/07/2002  
TIME: 15:16:25

Input Set : N:\CrF3\RULE60\09981947.raw  
Output Set: N:\CRF3\02072002\I981947.raw

L:20 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:21 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:598 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8  
L:608 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9  
L:619 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:625 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:629 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:633 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:649 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:651 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:677 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10  
L:679 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10